Interpolatory Nonlinear and Non-Separable Multi-scale Representation: Application to Image Compression

Sylvain Meignen, Basarab Matei, Anastasia Zakharova

Interpolatory Nonlinear and Non-Separable Multi-scale Representation: Application to Image Compression

Basarab Mateï

LAGA Laboratory, Paris XIII University, France

Tel:0033-1-49-40-35-71

FAX:0033-4-48-26-35-68

E-mail: matei@math.univ-paris13.fr

Sylvain Meignen*

LJK Laboratory, University of Grenoble, France

Tel:0033-4-76-51-43-95

FAX:0033-4-76-63-12-63

E-mail: sylvain.meignen@imag.fr

Anastasia Zakharova

LJK Laboratory, University of Grenoble, France

Tel:0033-4-76-51-43-95

FAX:0033-4-76-63-12-63

Email: anastasia.zakharova@gmail.com

Abstract- In this paper, we introduce the notion of nonlinear and non-separable multi-scale representation. We show how it can be derived from nonlinear and non-separable subdivision schemes associated to a non-diagonal dilation matrix. We focus on nonlinear multi-scale decomposition where the dilation matrix is either the quincunx or the hexagonal matrix. We then detail the encoding and decoding algorithm of the representation and, in particular, how the EZW (Embedded Zero-tree Wavelet) algorithm adapts in that context. Numerical experiments on image compression conclude the paper.

Keywords- Nonlinear multi-scale representation, image compression.

EDICS Category:TEC-MRS

I. INTRODUCTION

Due to the hierarchical structure of visual information, the multi-scale representations are widely applied in image processing [2][1][3]. Images are bidimensional complex objects made of homogeneous regions separated by smooth singularities. It is then natural to process them using techniques which are intrinsically bidimensional and are adapted to the information contained in an image. The multi-scale representations based on the tensor product of wavelet bases are not well adapted to process efficiently an image near the singularities. In order to have a better treatment of singularities, A. Harten in [5][6] introduced a general framework to multi-scale data representations. The idea is to associate to any function v, a set of sequences $\mathcal{M}v := (v^0, d^0, d^1, d^2, \ldots)$, where the sequence $v^0 := (v^0_k)_{k\in\mathbb{Z}^d}$ is the coarsest approximation of v and the sequences $d^j := (d^j_k)_{k\in\mathbb{Z}^d}$, with $j \geq 0$, are additional detail coefficients which represent the fluctuations of v between two successive levels of resolution.

There exist many ways to build such nonlinear representations. In most papers, the multi-scale structure is associated to dyadic levels of resolution [5][2]. In [4], the quincunx matrix is used to define the levels, and then a multi-scale representation is proposed. In the present paper, we build multi-scale representations either based on quincunx or hexagonal dilation matrices. The general philosophy of the approach, is to predict at a fine level using only coarser levels and to memorize prediction errors as detail coefficients. The approach we propose is therefore well adapted to progressive transmission of data.

We first recall some definitions about nonlinear subdivision operators in a non-separable framework and we will explain how this entails the convergence of multi-scale representations associated to such

operators. We will then give applications to interpolatory multi-scale representations corresponding to quincunx or hexagonal dilation matrices. In particular, we will study affine interpolatory scheme and also interpolation using polynomial of total degree two. The whole procedure from the encoding of the representation, to the compression step and, finally, to the decoding step will be detailed. In particular, we will show how the EZW (embedded zero-tree wavelets) adapts in our context. Some illustrations on image compression conclude the paper.

II. Notations and Multi-Resolution Analysis Definition

A. Notations

Let $\ell(\mathbb{Z}^d)$ be the space of all sequences indexed by $\mathbb{Z}^d$. The subspace of bounded sequences is denoted by $\ell^\infty(\mathbb{Z}^d)$ and $\|u\|_{\ell^\infty(\mathbb{Z}^d)}$ is the supremum of $\{|u_k| : k \in \mathbb{Z}^d\}$. We denote $\ell^0(\mathbb{Z}^d)$ the subspace of all sequences with finite support (i.e. the number of non-zero components of a sequence is finite). As usual, let $\ell^p(\mathbb{Z}^d)$ be the Banach space of sequences u on $\mathbb{Z}^d$ such that $\|u\|_{\ell^p(\mathbb{Z}^d)} < \infty$, where

$$\|u\|_{\ell^p(\mathbb{Z}^d)} := \left(\sum_{k\in\mathbb{Z}^d} |u_k|^p\right)^{\frac{1}{p}} \text{ for } 1 \le p < \infty.$$

For any $w \in \ell^p(\mathbb{Z}^d)^d$ we will denote $\|w\|_{\ell^p(\mathbb{Z}^d)^d}$, the supremum of the ℓ^p-norm of the components of w. As in the discrete case, we denote by $L^p(\mathbb{R}^d)$ the space of all measurable functions f such that $\|f\|_{L^p(\mathbb{R}^d)} < \infty$, where

$$\|f\|_{L^p(\mathbb{R}^d)} := \left(\int_{\mathbb{R}^d} |f(x)|^p dx\right)^{\frac{1}{p}} \text{ for } 1 \le p < \infty$$

and $\|f\|_{L^\infty(\mathbb{R}^d)}$ is the essential supremum of $|f|$ on $\mathbb{R}^d$. A matrix M is called a dilation matrix if it has integer entries and if $\lim_{n\to\infty} M^{-n} = 0$. In the following, the invertible dilation matrix is always denoted by M and m stands for $|\det(M)|$. Finally, for two positive quantities A and B depending on a set of parameters, the relation $A \lesssim B$ implies the existence of a positive constant C, independent of the parameters, such that $A \le CB$. Also $A \sim B$ means $A \lesssim B$ and $B \lesssim A$.

B. Multi-Resolution Analysis

To begin with, let V be a d-dimensional Hilbertian space. From here on, M will denote a dilation matrix. We recall the concept of multi-resolution analysis associated to the space V.

Definition 1. *A multi-resolution analysis of V is a sequence $(V_j)_{j\in\mathbb{Z}^d}$ of closed subspaces embedded in V satisfying the following properties:*

1) *The subspaces are embedded* $V_j \subset V_{j+1}$
2) $f \in V_j$ *if and only if* $f(M.) \in V_{j+1}$
3) $\overline{\cup_{j\in\mathbb{Z}} V_j} = V$.
4) $\cap_{j\in\mathbb{Z}} V_j = \{0\}$
5) *We assume the existence of a compactly supported function* $\varphi \in V_0$ *such that the family of their translates* $\{\varphi(\cdot - k)\}_{k\in\mathbb{Z}^d}$ *forms a Riesz basis for* V_0.

From $V_0 \subset V_1$, it follows that the function φ, called scaling function, satisfies:

$$\varphi(x) = \sum_{n\in\mathbb{Z}^d} g_n \varphi(Mx - n), \text{ with } \sum_n g_n = m. \tag{1}$$

To build the approximation of a given function v at level j, we adopt the biorthogonal point of view that is we assume the existence of a *dual* function $\tilde{\varphi}$ with compact support satisfying:

$$\tilde{\varphi}(x) = \sum_{k\in\mathbb{Z}^d} h_k \tilde{\varphi}(Mx - k), \text{ with } \sum_k h_k = m, \tag{2}$$

and such that the following duality property holds $< \tilde{\varphi}(x - n), \varphi(x - k) >= \delta_{n,k}$, where $\delta_{n,k}$ denotes the Kronecker symbol and $< ., . >$ the Euclidean inner product. The approximation at level j can be obtained by projection of v on V_j as follows:

$$v_j = \sum_{n\in\mathbb{Z}^d} v_n^j \varphi(M^j x - n) \tag{3}$$

where $v_k^j = m^j < v, \tilde{\varphi}(M^j x - k) >$. Note that when $\tilde{\varphi}$ is fixed, φ is a priori not unique which leads to potentially different v_j.

III. Multi-scale Representation Using Dilation Matrix M

Let $(\Gamma^j)_{j=0,\cdots,J}$ be the set of embedded grids with $\Gamma^j = \{M^{-j}k, k \in \mathbb{Z}^d\}$, J corresponding to the finest level of resolution. We now define the *projection* and *prediction* operators on these grids.

A. Projection and Prediction Operators

Let $(\Gamma^j)_{j=0,\cdots,J}$ be the multiresolution structure defined above and then $v^J = (v_k^J)_{k\in\mathbb{Z}^d}$ the data at the finest level J. The value v_k^J is associated to the location $M^{-J}k$ on the grid Γ^J. In order to build the multi-scale representation of v^J, we assume the existence of two discrete interscale operators:

1) the *projection* P_{j-1}^j operator acting from fine to coarse level. If $\hat{v}^j$ is an approximation of v^j on Γ^j , we define $v^{j-1} = P_{j-1}^j \hat{v}^j$. This operator is always assumed to be *linear*. In our context, $\tilde{\varphi}$

defined in (2) fixes the projection operator as follows:

$$v_k^{j-1} := (m)^{-1} \sum_{n \in \Gamma^j} h_{n-Mk} \hat{v}_n^j. \tag{4}$$

2) the *prediction* P_j^{j-1} operator acting from coarse to fine level. This operator computes $\hat{v}^j = P_j^{j-1} v^{j-1}$ and may be *nonlinear*.

We further assume that these two operators satisfy the *consistency* property: the projection of v^j on Γ^{j-1} coincides with v^{j-1}, i.e.

$$P_{j-1}^j P_j^{j-1} = I. \tag{5}$$

Note that the prediction error $e^j = v^j - \hat{v}^j$, is not arbitrary data on Γ^j. Indeed, from (5), one has

$$P_{j-1}^j e^j = P_{j-1}^j v^j - P_{j-1}^j \hat{v}^j = v^{j-1} - v^{j-1} = 0.$$

Hence, $e^j \in \mathrm{K}er(P_{j-1}^j)$. It follows that we may write this error in a non-redundant way using a basis of $\mathrm{K}er(P_{j-1}^j)$. The coefficients of the errors in this kernel gives the detail coefficients d^{j-1}. Consequently, the data v^j is completely equivalent to (v^{j-1}, d^{j-1}). Iterating this process from the initial data v^J, we obtain its *nonlinear multi-scale representation*

$$\mathcal{M} v^J = (v^0, d^0, \cdots, d^{J-1}).$$

B. Interpolatory Multi-scale Representations

We now introduce interpolatory multi-scale representations. In that context, the data at level j is:

$$v_k^j = v(M^{-j}k),$$

for any k in $\mathbb{Z}^d$. From this, we directly obtain the form of the projection operator (P_{j-1}^j):

$$v_k^{j-1} = v(M^{-j+1}k) = v_{Mk}^j$$

In that context, the principle of the multi-scale representation is based on the prediction of $v_{Mk+\varepsilon}^j$ for $\varepsilon \neq 0$ using the values of v^{j-1} on Γ_0^{j-1}. Knowing the prediction operator P_j^{j-1}, we then replace the value of $v_{Mk+\varepsilon}^j$, $\varepsilon \neq 0$ by the error: $e_{Mk+\epsilon}^j = v_{Mk+\epsilon}^j - (P_{j-1}^j v^{j-1})_{Mk+\epsilon}$. The multi-scale representation can then directly be written as $\mathcal{M} v^J = (v^0, d^0, \cdots, d^{J-1})$, where $d^{j-1} = \left\{ e_{Mk+\varepsilon}^j, \varepsilon \neq 0, k \in \mathbb{Z}^d \right\}$.

IV. Theoretical Results on Nonlinear Multi-Scale Representations

A. Definition of the Prediction Operator

The definition of the multi-scale representation is related to that of P_{j-1}^{j}, the *prediction* operator. In the present paper, we allow the definition and the use of *quasi-linear* prediction operator defined by:

Definition 1. *A quasi-linear prediction operator is a function which associates to each $w \in \ell^\infty(\mathbb{Z}^d)$ a linear operator $S(w)$ defined by*

$$(S(w)u)_k := \sum_{\{l\in\mathbb{Z}^d, \|k-Ml\|_\infty<K\}} a_{k-Ml}(w)u_l,$$

for any $u \in \ell^\infty(\mathbb{Z}^d)$ where $\|\cdot\|_\infty$ denote the sup norm in $\mathbb{Z}^d$ and

$$|a_{k-Ml}(w)| < C \quad \forall w.$$

The constants K and C are independent of the data w.

Note that the recursive action of the quasi-linear prediction operator $Su := S(u)u$ on the initial data $u = u^0$, defines a nonlinear subdivision scheme:

$$u^j := Su^{j-1} = S(u^{j-1})u^{j-1} = \ldots = S(u^{j-1})\ldots S(u^0)u^0 = S^j u^0.$$

We assume that the general form of the prediction operator for the given data v^{j-1} is then given by:

$$\hat{v}_k^j = (P_j^{j-1}v^{j-1})_k = (S(v^{j-1})v^{j-1})_k = \sum_{l\in\mathbb{Z}^d} a_{k-Ml}(v^{j-1})v_l^{j-1}.$$

The consistency property imposes that $a_{k,l}(v^{j-1})$ satisfies

$$\sum_{k\in\mathbb{Z}^d} a_{k-Mp}(v^{j-1})h_{k-Mi} = m\delta_{p-i}. \tag{6}$$

In an interpolatory framework, $h = m\delta_{k,0}$ and (6) amounts to $a_{Mk}(v^{j-1}) = \delta_{k,0}$.

B. Definitions of Schemes for the Differences and of Joint Spectral Radius

We say that the quasi-linear subdivision operator S reproduces the constants when:

$$\sum_{p\in\mathbb{Z}^d} a_{k-Mp}(u) = 1 \ \forall k \in \mathbb{Z}^d \text{ and } \forall u \in \ell^\infty(\mathbb{Z}^d).$$

Then, the following result holds:

Proposition 1. *Let S be a quasi-linear prediction operator reproducing the constants and assume $u^j = Su^{j-1}$ defined in (6). Then there exists a local and bounded operator S_1 such that $\nabla u^j := S_1(u^{j-1})\nabla u^{j-1}$. where $\nabla u_k = (u_k - u_{k-e_1}, \cdots, u_k - u_{k-e_d})$.*

PROOF: Consider

$$\begin{aligned} u^j_{k+e_i} - u^j_k &= \sum_{p \in V(k+e_i) \bigcup V(k)} (a_{k+e_i-Mp}(u^{j-1}) - a_{k-Mp}(u^{j-1}))u^{j-1}_p \\ &= \sum_{p \in V(k+e_i) \bigcup V(k)} \alpha_{k-Mp}(u^{j-1})u^{j-1}_p, \end{aligned}$$

where $V(k) = \{p, \|k - Mp\|_\infty < K\}$. Note that $\sum_{p \in V(k+e_i) \bigcup V(k)} \alpha_{k-Mp}(u^{j-1}) = 0$, we then deduce that, since $\left\{\nabla_l \delta_{n,\beta}, n \in V(k+e_i) \bigcup V(k), \beta \in \mathbb{Z}^d, l = 1, \cdots, d\right\}$ (with $\nabla_l u_k = u_k - u_{k-e_l}$) spans $\left(\alpha_{k-Mp}(u^{j-1})\right)_{p \in V(k+e_i) \bigcup V(k)}$ [7], $u^j_{k+e_i} - u^j_k = \sum_{\beta \in \mathbb{Z}^d} \sum_{p \in V(k+e_i) \bigcup V(k)} \sum_{l=1}^{d} c_{p-\beta,l} \nabla_l u^{j-1}_p$, where (c_k) is a finite sequence. Computing the differences for other directions e_i, we obtain the desired result ■.

C. Convergence Theorem for the Multi-scale Representation

In this section, we state a convergence result for the multi-scale representation when the matrix M is isotropic which corresponds to the following definition:

Definition 2. *We say that a matrix M is isotropic if it is similar to the diagonal matrix diag$(\sigma_1, \ldots, \sigma_d)$, i.e. there exists an invertible matrix Λ such that*

$$M = \Lambda^{-1} \mathit{diag}(\sigma_1, \ldots, \sigma_d)\Lambda,$$

with $|\sigma_1| = \ldots = |\sigma_d|$ being the eigenvalues of matrix M.

For an isotropic matrix holds $|\sigma_1| = \ldots = |\sigma_d| = \sigma = m^{\frac{1}{d}}$. Moreover, for any given norm in $\mathbb{R}^d$ there exist constants C_1, C_2 such that for any integer n and for any $v \in \mathbb{R}^d$

$$C_1 \sigma^n \|v\| \leq \|M^n v\| \leq C_2 \sigma^n \|v\|.$$

The convergence theorems we now state involve Besov spaces. Let us briefly recall the definition of such spaces. First consider the modulus of smoothness of order $N \in \mathbb{Z}$ in $L^p(\mathbb{Z}^d)$ of a function v, given by

$$\omega_N(v,t)_{L^p} = \sup_{h \in \mathbb{Z}^d : \|h\|_2 \leq t} \|\nabla^N_h v\|_{L^p(\mathbb{R}^d)},$$

where $\|.\|_2$ is the Euclidian norm and where ∇_h^N is the finite differences operator of order N in the direction h:

$$\nabla_h^N v(x) := \sum_{\beta=0}^{N} (-1)^\beta C_N^\beta v(x - \beta h)$$

We now introduce the Besov space $B^s_{p,q}(\mathbb{R}^d)$. Let N be any integer, $N > s$. For $p, q > 0$ and $s > 0$, the space $B^s_{p,q}(\mathbb{R}^d)$ contains functions $v \in L^p(\mathbb{R}^d)$ such that $(2^{js}\omega_N(v, 2^{-j})_{L^p})_{j\geq 0} \in \ell^q(\mathbb{Z}^d)$. The norm in $B^s_{p,q}(\mathbb{R}^d)$ is then given by $\|v\|_{B^s_{p,q}(\mathbb{R}^d)} := \|v\|_{L^p(\mathbb{R}^d)} + \|(2^{js}\omega_N(v, 2^{-j})_{L^p})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}$. If we consider an isotropic dilation matrix M, we define from v the set v_j ,$j \geq 0$ converging to v following (3) and then the nonlinear multi-scale representation. Then, we have the following theorem:

Theorem 1. *If the prediction operator reproduces the constants, then for all* $0 < s < 1$*, one has the inequality*

$$\|v_\cdot^0\|_{\ell^\infty(\mathbb{Z}^d)} + \|(\sigma^{sj}\|d_\cdot^{j-1}\|_{\ell^\infty(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|v\|_{B^s_{\infty,q}(\mathbb{R}^d)}.$$

PROOF: First, we show that the consistency property implies that $\|e^j\|_{\ell^\infty(\mathbb{Z}^d)} \lesssim \|\nabla v^j\|_{\ell^\infty(\mathbb{Z}^d)^d}$. We recall that

$$e_k^j := v_k^j - \hat{v}_k^j = v_k^j - \sum_{\|k-Ml\|_\infty \leq K} a_{k-Ml}(v^{j-1}) v_l^{j-1}.$$

Using (4), we write the prediction error in the form

$$\begin{aligned} e_k^j &= v_k^j - m^{-1} \sum_{l:\|k-Ml\|_\infty \leq K} a_{k-Ml}(v^{j-1}) \sum_{p:\|p-Ml\|_\infty \leq P} \tilde{h}_{p-Ml} v_p^j \\ &= v_k^j - m^{-1} \sum_{p:\|k-p\|_\infty \leq K+P} v_p^j \sum_{l:\|k-Ml\|_\infty \leq K} a_{k-Ml}(v^{j-1}) \tilde{h}_{p-Ml} \\ &= \sum_{p\in F(k)} b_{k,p}(v^{j-1}) v_p^j, \end{aligned}$$

where $F(k) = \{p : \|p - k\|_\infty \leq P + K\}$ is a finite set for any given k. Let us define, for each $k \in \mathbb{Z}^d$, a vector $b_k(w) = (b_{k,n}(w))_{n\in F(k)}$. By hypothesis, $e^j = 0$ if $v_k^j = k$, since the prediction operator reproduces the constants. Since, $\{(\nabla_i \delta_{n-l})_{n\in F(k)}, l \in \mathbb{Z}^d, i = 1, \cdots, d\}$ spans the orthogonal to the family $(l_i^{q_i})_{l\in F(k), i=1,\cdots,d, q_i\leq 1}$, by denoting $c_l(w)$ the coordinates of $b_k(w)$ in this basis and after simple computations we obtain that

$$e_k^j = \sum_{n\in F(k)} \sum_{i=1}^{d} c_{n,i}(v^{j-1}) \nabla_i v_n^j.$$

This proves that $\|e^j\|_{\ell^\infty(\mathbb{Z}^d)} \lesssim \|\nabla v^j\|_{\ell^\infty(\mathbb{Z}^d)}$ since the sequence c is finitely supported. We get from this that

$$\|v^0\|_{\ell^\infty(\mathbb{Z}^d)} + \|(\sigma^{sj}\|d_\cdot^{j-1}\|_{\ell^\infty(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \|(\sigma^{sj}\|\nabla v_\cdot^j\|_{(\ell^\infty(\mathbb{Z}^d))^d})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}. \quad (7)$$

Then remark that $\|v^0\|_{\ell^\infty(\mathbb{Z}^d)} = \|v(.)\|_{\ell^\infty(\mathbb{Z}^d)} \lesssim \|v\|_{L^\infty(\mathbb{R}^d)}$ and also, since the matrix is isotropic, we can show that $\|\nabla v^j\|_{\ell^\infty(\mathbb{Z}^d)^d} \lesssim \omega_1(v, C\sigma^{-j}t) \lesssim \omega_1(v, \sigma^{-j}t)$, the last inequality being a property of the modulus of smoothness. Finally, replacing in equation (7), we get that:

$$\|v^0\|_{\ell^\infty(\mathbb{Z}^d)} + \|(\sigma^{sj}\|d_\cdot^{j-1}\|_{\ell^\infty(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|v\|_{B^s_{\infty,q}(\mathbb{R}^d)} \blacksquare.$$

Note that the joint spectral radius is not involved in the above inequality but is of crucial importance for the inverse theorem which we now state. We study the reconstruction process of a function v_r, where r stands for "reconstruction", from its coefficients $v^{0,r}$ and $d^{j,r}$. In the case of nonlinear representations, it is still possible to study the reconstruction algorithm of the function v_r from its coefficients by iteratively using the reconstruction step

$$v^{j,r} = P_j^{j-1}v^{j-1,r} + e^j = P_j^{j-1}v^{j-1,r} + Ed^{j-1,r}.$$

In that context, the function v_r is the limit (when it exists) of $v_{j,r}$ defined by:

$$v_{j,r}(x) = \sum_{k\in\mathbb{Z}^d} v_k^{j,r}\varphi(M^j x - k)$$

where φ was defined in (1). The following theorem gives a sufficient condition for the limit function to exist in Besov space:

Theorem 2. *Let S be a quasi-linear prediction operator reproducing the constants. If $\rho_\infty(S_1) < \sigma^{-s}$ for some $s > 0$ and if $(v^{0,r}, d^{0,r}, d^{1,r}, \ldots)$ are such that*

$$\|v_\cdot^{0,r}\|_{\ell^\infty(\mathbb{Z}^d)} + \|(\sigma^{sj}\|d_\cdot^{j,r}\|_{\ell^\infty(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} < \infty,$$

and assume that $\sum_{k\in\mathbb{Z}^d} \varphi(x-k) = 1$, where ϕ satisfies (1), then the function v belongs to $B^s_{\infty,q}(\mathbb{R}^d)$ and

$$\|v\|_{B^s_{\infty,q}} \lesssim \|v^0\|_{\ell^\infty(\mathbb{Z}^d)} + \|(\sigma^{sj}\|d_\cdot^j\|_{\ell^\infty(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}.$$

The proof of the inverse theorem is much more involved and details are available in [10]. Note that the inverse theorem requires that the joint spectral radius of the difference operator be smaller than 1. We will design in the following sections examples of prediction operators that satisfy this property.

D. Stability of Nonlinear Multi-scale representations

The stability of nonlinear multi-scale decomposition is strongly related to that of the associated prediction operator. We recall the important notion of the stability of quasi-linear prediction operator [2]:

Definition 3. *A quasi-linear prediction operator is stable if for every v, w in $\ell^\infty(\mathbb{Z}^d)$:*

$$\|S(v) - S(w)\|_{\ell^\infty(\mathbb{Z}^d)} \leq C\|v - w\|_{\ell^\infty(\mathbb{Z}^d)}$$

where C depends in a non-decreasing way on $\|v\|_{\ell^\infty(\mathbb{Z}^d)}$.

The stability of the quasi-linear prediction operator is essential for the stability of the multi-scale representation $\mathcal{M}v^J$. We will tell when this property holds for the prediction operators we will consider, but the study of the stability of the associated multi-scale representation is beyond the scope of the present article.

V. Bidimensional Interpolatory Quasi-Linear Prediction Operators

The nonlinear representations we will study are essentially based on edge non-oscillatory (ENO) approach in an interpolatory framework. ENO methods consist in computing a cost function denoted by $C^j(k)$ that determines the best prediction operator among a group of potential ones. Each of these prediction operators are associated to interpolation polynomials on different stencils, the cost function $C^j(k)$ being a function computed on the stencils. An arbitrarily small change at the round off level would be sufficient to change the stencil chosen for the prediction and thus the prediction operator. For that reason, the ENO scheme is not stable [10]. Such a drawback can be avoided using weighted-ENO (WENO) interpolation which provides a smooth transition between prediction operators. The WENO formulation is based on a convex combination of potential prediction operators given by the ENO method, that is:

$$\hat{v}_k^j := \sum_{r=0}^{m-1} \alpha_r \hat{v}_k^{j,r}$$

with $\alpha_r \geq 0$ and $\sum_{r=0}^{m-1} \alpha_r = 1$. A possible form for the weights is given by $\alpha_r := \frac{a_r}{\sum_{i=0}^{m-1} a_i}$ for $r = 0, \ldots, m-1$ where $a_r := \frac{1}{\varepsilon + C^j(r)}$, where $C^j(r)$ is the cost function of the corresponding rule. We will see, on examples that the WENO prediction is continuously dependent on the data.

In the following, we build nonlinear multi-scale decompositions using the quincunx or the hexagonal dilation matrix which are particular cases of isotropic matrices. A similar approach using the quincunx

matrix was proposed in [4] in a lifting scheme framework, we will emphasize the differences with our method when necessary. In one dimension and in a dyadic scales framework, the lifting scheme is made of an update step at even locations and a prediction step at odd locations but, when one considers interpolatory scheme to predict, as we will do, no update is required. Since we consider an interpolatory scheme, $h = m\delta_0$ and the consistency property leads to $a_{Mk} = \delta_{k,0}$. Consequently, we only define $a_{Mk+\varepsilon}$ when ε is a non-zero coset vector of M.

A. Nonlinear Affine Prediction Using the Quincunx Matrix

The quincunx matrix is defined by

$$M = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix},$$

whose coset vectors are $\varepsilon_0 = (0,0)^T$ and $\varepsilon_1 = (0,1)^T$. Note that $a_{0,0} = 1$ since the scheme is interpolatory (this corresponds to $v^j_{Mk} = v^{j-1}_k$) and since the nonlinear prediction operators reproduce the constants we shall have $\sum_i a_{Mi+\epsilon} = 1$ for all coset vectors ϵ. We consider prediction operators based on interpolation by polynomials of degree 1 (i.e. $a+bx+cy$) on Γ^{j-1}, leading the following four potential stencils which in turn entails the following prediction rules:

$$\hat{v}^{j,1}_{Mk+\varepsilon_1} = \tfrac{1}{2}(v^{j-1}_k + v^{j-1}_{k+e_1+e_2}) \tag{8}$$

$$\hat{v}^{j,2}_{Mk+\varepsilon_1} = \tfrac{1}{2}(v^{j-1}_{k+e_1} + v^{j-1}_{k+e_2}) \tag{9}$$

Now, as $M^2 = 2Id$, after double iteration these subdivision schemes could be associated with a limit function of the kind:

$$\Psi(x) = \sum_p \sum_k (a_{\uparrow M})_k a_{p-k} \Psi(2x-p)$$

where

$$(a_{\uparrow M})_p = \begin{cases} a_k \text{ if } p = Mk \\ 0 \text{ otherwise,} \end{cases}$$

and, therefore, we can draw a comparison between the different cases in terms of limit functions. On Figure 1.C, we display the limit function corresponding to the separable case (i.e. $\Psi(x) = \prod_{i=1}^{2} \max(0, 1-|x_i|)$) where the predictions (8) and (9) are applied successively, while we display on Figure 1. A (resp. B), the limit function corresponding to prediction (8) (resp. (9)). To relate this to the general biorthogonal framework of section II, we shall say that the limit function corresponding to the separable case is the function φ and the other predictions define other limit functions that are still orthogonal to $\tilde{\varphi}$.

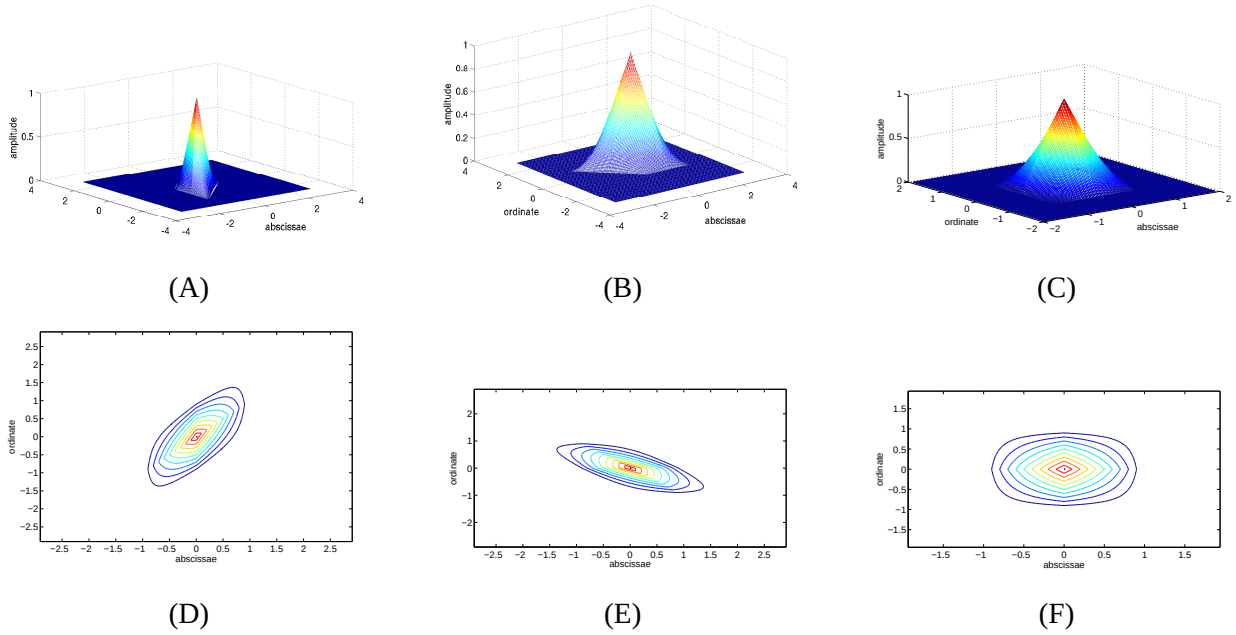


Fig. 1. (A): the limit function associated with prediction (8),(B): the limit function associated with prediction (9); (C): the limit function corresponding to the separable case; (D), (E)and (F): the corresponding contour plots.

We now pass on to ENO prediction, which consists in choosing between the two prediction rules (8) and (9). In [4], the choice of stencils corresponds to:

$$\min\left(|v^j_{Mk+\varepsilon_1}-\frac{1}{2}(v^{j-1}_{k+e_1}+v^{j-1}_{k+e_2})|,|v^j_{Mk+\varepsilon_1}-\frac{1}{2}(v^{j-1}_{k}+v^{j-1}_{k+e_1+e_2})|\right).$$

The above cost function assumes that the true values at level j are known before prediction which is not relevant in a progressive data transmission context. To avoid this, we define the cost function as follows:

$$C^j(k)=\min(|v^{j-1}_{k+e_1}-v^{j-1}_{k+e_2}|,|v^{j-1}_{k}-v^{j-1}_{k+e_1+e_2}|).$$

When the minimum of $C^j(k)$ corresponds to the first (resp. second) argument the prediction (9) (resp. (8)) is used. One motivation for the choice of such a cost function is the following argument: when an edge intersect the cell Q^{j-1}_k delimited by the points $M^{-j+1}\{k,k+e_1,k+e_2,k+e_1+e_2\}$ of the grid Γ^{j-1}, several cases may happen:

1) either the edge intersect $[M^{-j+1}k,M^{-j+1}(k+e_1+e_2)]$ and $[M^{-j+1}(k+e_1),M^{-j+1}(k+e_2)]$ in which case no direction is favored.
2) or the edge intersect $[M^{-j+1}k,M^{-j+1}(k+e_1+e_2)]$ or $[M^{-j+1}(k+e_1),M^{-j+1}(k+e_2)]$, in which case the prediction operator should favor the direction which is not intersected by the edge (this is what the cost function does).

When Q_k^{j-1} is not intersected by an edge, the gain between choosing one direction or the other is very small [4]. Therefore, we will only apply this procedure for stencil selection only when:

$$\begin{aligned}
&\underset{k'=k,k+e_1,k-e_1}{\text{argmin}} \quad \Big(|v_{k'}^{j-1}-v_{k'+e_1}^{j-1}|+|v_{k'}^{j-1}-v_{k'+e_2}^{j-1}|+|v_{k'+e_2}^{j-1}-v_{k'+e_1+e_2}^{j-1}| \\
&\qquad +|v_{k'+e_1}^{j-1}-v_{k'+e_1+e_2}^{j-1}|+|v_{k'}^{j-1}-v_{k'+e_1+e_2}^{j-1}| \quad +|v_{k'+e_1}^{j-1}-v_{k'+e_2}^{j-1}|\Big)=1 \text{ or} \\
&\underset{k'=k,k+e_2,k-e_2}{\text{argmin}} \quad \Big(|v_{k'}^{j-1}-v_{k'+e_1}^{j-1}|+|v_{k'}^{j-1}-v_{k'+e_2}^{j-1}|+|v_{k'+e_2}^{j-1}-v_{k'+e_1+e_2}^{j-1}| \\
&\qquad +|v_{k'+e_1}^{j-1}-v_{k'+e_1+e_2}^{j-1}|+|v_{k'}^{j-1}-v_{k'+e_1+e_2}^{j-1}|+|v_{k'+e_1}^{j-1}-v_{k'+e_2}^{j-1}|\Big)=1 \qquad (10)
\end{aligned}$$

which corresponds to the case where the average first order differences are locally maximum in the vertical or horizontal direction. We will also test another possible choice to detect the cells where the nonlinear prediction is potentially interesting which is the following:

$$\begin{aligned}
&\underset{k'=k,k+e_1+e_2,k-e_1-e_2}{\text{argmin}} \quad (|v_{k'}^{j-1}-v_{k'+e_1+e_2}^{j-1}|)=1 \text{ or} \\
&\underset{k'=k,k+e_1-e_2,k-e_1+e_2}{\text{argmin}} \quad (|v_{k'+e_1}^{j-1}-v_{k'+e_2}^{j-1}|)=1 \qquad (11)
\end{aligned}$$

which corresponds to the case where the first order differences are locally maximum is the direction of prediction. When the cell does not satisfy this property, we will apply the linear prediction method.

As far as the WENO approach is concerned, no cost function is needed since we have:

$$\tilde{v}_{Mk+\varepsilon_1}^j=\frac{a_1}{2(a_1+a_2)}(v_{k+e_1}^{j-1}+v_{k+e_2}^{j-1})+\frac{a_2}{2(a_1+a_2)}(v_k^{j-1}+v_{k+e_1+e_2}^{j-1})$$

with

$$a_1=\frac{1}{\epsilon+|v_{k+e_1}^{j-1}-v_{k+e_2}^{j-1}|}, a_2=\frac{1}{\epsilon+|v_k^{j-1}-v_{k+e_1+e_2}^{j-1}|}. \qquad (12)$$

The ENO and WENO prediction are such that $\rho_\infty(S_1)<1$ since they satisfy the following property:

Proposition 2.

- *when $k=Mk'$, we can show that:*

$$\begin{aligned}
\|v_{Mk+\varepsilon_1}^{j,1}-v_{Mk}^j\|_{l^\infty(\mathbb{Z}^2)} &\leq \frac{1}{2}\|\nabla v_\cdot^{j-2}\|_{(l^\infty(\mathbb{Z}^2))^2} \\
\|v_{Mk+\varepsilon_1}^{j,2}-v_{Mk}^j\|_{l^\infty(\mathbb{Z}^2)} &\leq \|\nabla v_\cdot^{j-2}\|_{(l^\infty(\mathbb{Z}^2))^2}
\end{aligned}$$

- *when $k=Mk'+\varepsilon_1$, we can show that:*

$$\begin{aligned}
\|v_{Mk+\varepsilon_1}^{j,2}-v_{Mk}^j\|_{l^\infty(\mathbb{Z}^2)} &\leq \frac{1}{2}\|\nabla v_\cdot^{j-2}\|_{(l^\infty(\mathbb{Z}^2))^2} \\
\|v_{Mk+\varepsilon_1}^{j,1}-v_{Mk}^j\|_{l^\infty(\mathbb{Z}^2)} &\leq \|\nabla v_\cdot^{j-2}\|_{(l^\infty(\mathbb{Z}^2))^2}
\end{aligned}$$

The proof is available in Appendix B. The same result can be shown for other differences on the grid Γ^j, which proves that $\rho_\infty(S_1)<1$ and then enables us to apply Theorem 2.

B. Nonlinear Affine Prediction Using the Hexagonal Dilation Matrix

We now focus on the construction of nonlinear multi-scale decompositions using as dilation matrix the hexagonal matrix

$$M = \begin{pmatrix} 2 & 1 \\ 0 & -2 \end{pmatrix},$$

with coset vectors $\varepsilon_0 = (0,0)^T, \varepsilon_1 = (1,0)^T, \varepsilon_2 = (1,-1)^T, \varepsilon_3 = (2,-1)^T$. We compute the prediction of v^j for the different coset points using an affine interpolant of the neighboring points associated to values computed on the grid Γ^{j-1}. To do so, we use the following four different stencils on the grid Γ^{j-1}:

$$\begin{aligned} V_k^{j,1} &= M^{-j+1}\{k, k+e_1, k+e_2\}, \\ V_k^{j,2} &= M^{-j+1}\{k, k+e_2, k+e_1+e_2\}, \\ W_k^{j,1} &= M^{-j+1}\{k+e_1, k+e_2), k+e_1+e_2\}, \\ W_k^{2} &= M^{-j+1}\{k, k+e_1, k+e_1+e_2\}. \end{aligned}$$

We determine to which stencils each point of Γ^j belongs to, and we then define the prediction as its barycentric coordinates. Note, first, that the prediction rule at Mk and $Mk+\varepsilon_1$ is independent of the choice of the stencil, and we always have:

$$v_{Mk}^j = v_k^{j-1} \text{ and } v_{Mk+\varepsilon_1}^j = \frac{1}{2}v_k^{j-1} + \frac{1}{2}v_{k+e_1}^{j-1}. \tag{13}$$

When one considers the prediction for the coset vector ε_2, V_k^1 or V_k^2 can be used to predict leading respectively to:

$$\begin{aligned} v_{Mk+\varepsilon_2}^{j,1} &= \frac{1}{4}v_{k+e_1}^{j-1} + \frac{1}{2}v_{k+e_2}^{j-1} + \frac{1}{4}v_k^{j-1} \\ v_{Mk+\varepsilon_2}^{j,2} &= \frac{1}{2}v_k^{j-1} + \frac{1}{4}v_{k+e_2}^{j-1} + \frac{1}{4}v_{k+e_1+e_2}^{j-1}. \end{aligned} \tag{14}$$

When one considers the prediction rules for the coset vector ε_3, W_k^1 or W_k^2 can be used leading to the following two predictions:

$$\begin{aligned} v_{Mk+\varepsilon_3}^{j,1} &= \frac{1}{4}v_{k+e_2}^{j-1} + \frac{1}{4}v_{k+e_1+e_2}^{j-1} + \frac{1}{2}v_{k+e_1}^{j-1} \\ v_{Mk+\varepsilon_3}^{j,2} &= \frac{1}{4}v_k^{j-1} + \frac{1}{4}v_{k+e_1}^{j-1} + \frac{1}{2}v_{k+e_1+e_2}^{j-1} \end{aligned} \tag{15}$$

when the stencils W_k^1 and W_k^2 are used respectively.

This leads to four different linear prediction rules depending on the choice for the prediction operator for coset vector ε_2 and ε_3. The corresponding limit functions are shown on Figure 2. To relate this to the general biorthogonal framework of section II, the function φ would be associated to the predictions (14.1) and (15.1) while other prediction rules defines other limit function that are still biorthogonal to $\tilde{\varphi}$. The proposed nonlinear subdivision operator is such that $\rho_\infty(S_1) < 1$ since we have:

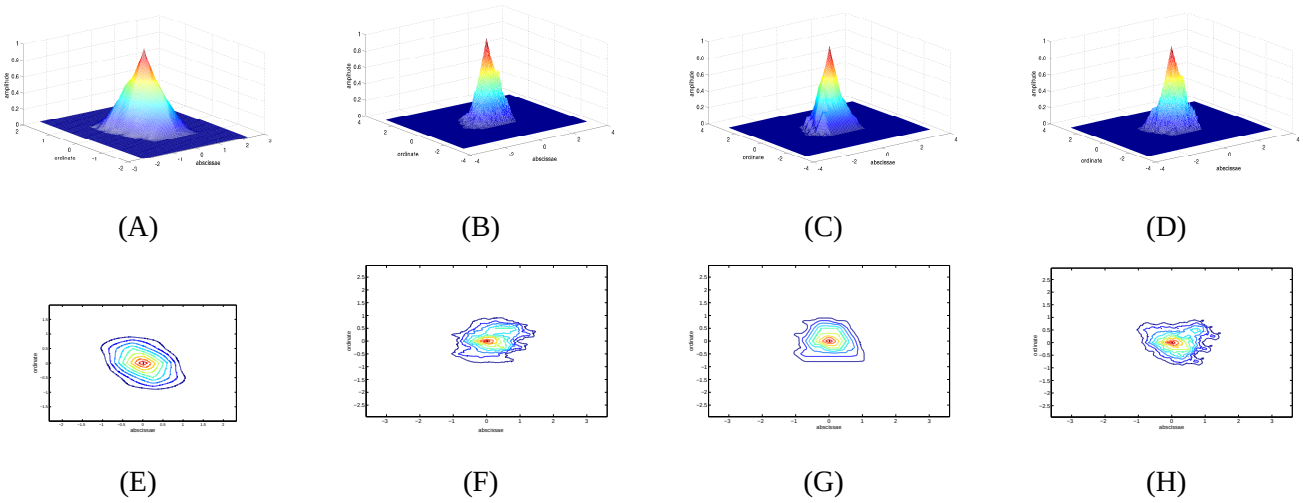


Fig. 2. (A): the limit function associated to predictions (14.1), (15.1); (B): the limit function associated to predictions (14.1), (15.2); (C): idem with the predictions (14.2), (15.1); (D): idem with prediction (14.2), (15.2) . The corresponding contour plots are depicted on figures (E),(F),(G) and (H) respectively.

Proposition 3. *The prediction defined by (13), (14), (15) satisfies:*

$$\|\nabla v^j_{M.+\epsilon_i}\|_{(l^\infty(\mathbb{Z}^2))^2} \leq \frac{3}{4}\|\nabla v^{j-1}\|_{(l^\infty(\mathbb{Z}^2))^2}$$

The proof is given in Appendix C.

The choice of stencil in the ENO method is determined by minimizing a certain cost function, which is different for ε_2 and ε_3 since these coset points belong to different stencils. We will consider the following cost function to predict at ε_2:

$$C_H^{j,\varepsilon_2,1}(k) = \min(|v^{j-1}_{k+e_1} - v^{j-1}_{k+e_2}| + |v^{j-1}_{k+e_1} - v^{j-1}_{k}|, |v^{j-1}_{k} - v^{j-1}_{k+e_1+e_2}| + |v^{j-1}_{k+e_1+e_2} - v^{j-1}_{k+e_1}|).$$

When the minimum corresponds to the first (resp. second) argument, the stencil V_k^1 (resp. V_k^2) is used. Similarly, when one considers the prediction at ε_3, we will compute:

$$C_H^{j,\varepsilon_3,1}(k) = \min(|v^{j-1}_{k+e_1} - v^{j-1}_{k+e_2}| + |v^{j-1}_{k+e_1+e_2} - v^{j-1}_{k+e_2}|, |v^{j-1}_{k} - v^{j-1}_{k+e_1}| + |v^{j-1}_{k+e_1+e_2} - v^{j-1}_{k}|).$$

When the minimum corresponds to the first (resp. second) argument, the stencil W_k^1 (resp. W_k^2) is used. The above cost function to determine which stencil to use will apply only on cells Q_k^{j-1} containing an edge. Such cells are determined in the same way as in the quincunx case finding local maxima of first order differences computed on the cell. When the hexagonal dilation matrix is used, to compute the cost function in the direction of prediction as in the quincunx case does not make sense since the direction of prediction does not exist strictly speaking. Therefore, to compute edge cells we will only consider formula (10).

As far as WENO prediction is concerned, we may write that:

$$\tilde{v}^j_{Mk+\varepsilon_2} = (\frac{\alpha_1^1}{4} + \frac{\alpha_2^1}{2})v_k^{j-1} + \frac{\alpha_1^1}{4}v_{k+e_1}^{j-1} + (\frac{\alpha_1^1}{2} + \frac{\alpha_2^1}{4})v_{k+e_2}^{j-1} + \frac{\alpha_2^1}{4}v_{k+e_1+e_2}^{j-1},$$
$$\tilde{v}^j_{Mk+\varepsilon_3} = (\frac{\alpha_1^2}{4} + \frac{\alpha_2^2}{4})v_{k+e_1}^{j-1} + \frac{\alpha_2^2}{2}v_k^{j-1} + (\frac{\alpha_1^2}{2} + \frac{\alpha_2^2}{4})v_{k+e_1+e_2}^{j-1} + \frac{\alpha_1^2}{4}v_{k+e_2}^{j-1},$$

with

$$a_1^1 = \frac{1}{\varepsilon + |v_{k+e_1}^{j-1} - v_{k+e_2}^{j-1}|}, a_2^1 = \frac{1}{\varepsilon + |v_k^{j-1} - v_{k+e_1+e_2}^{j-1}|},$$
$$a_1^2 = \frac{1}{\varepsilon + |v_{k+e_1}^{j-1} - v_{k+e_2}^{j-1}|}, a_2^2 = \frac{1}{\varepsilon + |v_k^{j-1} - v_{k+e_1+e_2}^{j-1}|}. \quad (16)$$

The WENO predictions satisfy the following stability property:

Theorem 3. *If a_r are defined by (12) for quincunx matrix and (16) for hexagonal matrix, the corresponding WENO subdivision schemes are stable.*

The proof is given in Appendix A.

C. Prediction Operators Using Higher Degree Polynomials

We now introduce prediction rules based on higher degree polynomials. Prediction Operators build using an affine interpolation suffers from the drawback that the prediction inside a given cell does not depend on neighboring cells. To consider higher degree polynomials enables to build more sophisticated prediction operators but results in spurious oscillations close to edges due to Gibbs phenomenon. In [8], the PPH scheme is introduced and aims at limitating oscillations close to the edges. It is derived from the one dimensional prediction operator defined as follows on a dyadic grid:

$$v_{2k+1}^j = \begin{cases} \frac{v_{k+1}^{j-1}+v^{j-1}k}{2} - \frac{1}{4}\frac{\nabla^2 v_k^{j-1}\nabla^2 v_{k+1}^{j-1}}{\nabla^2 v_k^{j-1}+\nabla^2 v_{k+1}^{j-1}} \text{ if } \nabla^2 v_k^{j-1}\nabla^2 v_{k+1}^{j-1} > 0 \\ \frac{v_{k+1}^{j-1}+v_k^{j-1}}{2} \text{ otherwise} \end{cases}$$

where $\nabla^2 v_k^{j-1}$ is the second order divided difference based on $x_{k-1}^{j-1}, x_k^{j-1}, x_{k+1}^{j-1}$. From this prediction operator, one can derive a bidimensional prediction operator using a tensor product approach. The convergence of the one dimensional subdivision scheme was shown in [8] and also the stability in [9]. The philosophy of the PPH approach is different from that we propose next in that the prediction operator is neither quasi-linear nor non-separable. Furthermore, the convergence of the bidimensional scheme derived by tensor product remains an open issue.

In what follows, we consider polynomials of degree 2 for prediction. These polynomials interpolates v on different stencils which we define below. The approach we propose next is valid both for the quincunx and the hexagonal dilation matrices. The prediction rules we consider are based on the following stencils defined on Γ^{j-1}:

$$\begin{aligned} V^1 &= M^{-j+1}\{k, k+e_1, k+e_2, k+e_1+e_2, k+2e_1, k+2e_2\} \\ V^2 &= M^{-j+1}\{k, k+e_1, k+e_2), k+e_1+e_2, k-e_1, k+2e_2\} \\ V^3 &= M^{-j+1}\{k, k+e_1, k+e_2, k+e_1+e_2, k+2e_1, k-e_2\} \\ V^4 &= M^{-j+1}\{k, k+e_1, k+e_2, k+e_1+e_2, k-e_1, k-e_2\}. \end{aligned} \tag{17}$$

As in the affine case, to detect edge-cells, we check property (10). Then, we use nonlinear affine prediction on edge-cells. Once we have dealt with edge-cells, we apply the following strategy on remaining cells:

- If a cell has a edge-cell as neighboring cell in the vertical or horizontal direction then the prediction is made using stencil selection we describe below,
- otherwise, we apply the stencil V^1 for the prediction.

Now, let us explain how we do stencil selection. For each stencil V^i, defined in (17), we consider the triangles made of neighboring points inside that stencil. This corresponds to 6 triangles. For instance for stencil V^1, the triangles are as follows: $M^{-j+1}\{k, k+e_1, k+e_2\}, M^{-j+1}\{k, k+e_1, k+e_1+e_2\}, M^{-j+1}\{k, k+e_2, k+e_1+e_2\}, M^{-j+1}\{k+e_1, k+e_2, k+e_1+e_2\}, M^{-j+1}\{k+e_1, k+2e_1, k+e_1+e_2)\}, M^{-j+1}\{k+e_2, k+e_1+e_2, k+2e_2\}$. The triangles are defined the same way for other stencils. We compute a cost function on each triangle as the sum of the first order differences along its edges (for instance, the cost function for the first triangle of V^1 is $|v_{k+e_1}^{j-1} - v_k^{j-1}| + |v_{k+e_2}^{j-1} - v_{k+e_1}^{j-1}| + |v_{k+e_2}^{j-1} - v_k^{j-1}|$). The cost function $C^{j,i}(k)$ associated to V^i is then the sum of the cost functions computed on all the triangles that make up V^i. For the prediction, we then use the stencil with minimal cost.

In the case where M is the quincunx matrix, the stencils V^i are used to predict $v_{Mk+\varepsilon_1}^j$, $\varepsilon_1 = (0,1)$ while when M is the hexagonal matrix, the stencils are used to predict $v_{Mk+\varepsilon_i}^j$, with $\varepsilon_1 = (1,0), \varepsilon_2 =$

$(1,-1), \varepsilon_3 = (2,-1)$. To prove the convergence of such schemes would involve the computation of all the differences in the canonical directions. However, at each coset point we have this time 4 potential predictions, so it would be prohibitive to detail the computation here. We are currently looking for a concise proof of the conjecture that these nonlinear operators satisfy the condition on the joint spectral radius (i.e. $\rho_\infty(S_1) < 1$). Once the prediction operator is chosen, we implement the multi-scale representation $\mathcal{M}v^J$. This is what we deal with next.

VI. Multi-scale Representation when $M^2 = \lambda Id$

Note first that since $M^2 = \lambda Id$, the finest resolution level J is considered even. Let $\Gamma^J = \{x^J_{k_1}, y^J_{k_2}\}_{k_1,k_2=0}^{N_J-1}, x^J_{k_1} = k_1 h_J, y^J_{k_2} = k_2 h_J, h_J = 2^{-J}h_0, N_J = 2^J N_0$ where N_0 is some integer and $h_0 = 1/N_0$. Since $M^2 = \lambda Id$, we obtain, for $k_1, k_2 = 0, ..., N_J/\lambda - 1$, $x^J_{\lambda k_1} = x^{J-2}_{k_1}$ and $y^J_{\lambda k_2} = y^{J-2}_{k_2}$. The connections between the levels J and $J-1$ or $J-1$ and $J-2$ are more elaborate.

A. Encoding and Decoding Algorithm

Let us detail the quincunx case, the hexagonal case can be dealt with the same way. The following encoding and decoding algorithms were originally proposed in [11], but we recall it for the sake of consistency.

We recall that the quincunx matrix satisfies $M^2 = 2Id$. For the first step, we have for $k_2 = 0, ..., N_J - 1$

$$(x^J_{2k_1}, y^J_{k_2}) = (x^{J-1}_{k_1}, y^{J-1}_{k_2}), \quad k_1 = 0, ..., N_J/2 - 1 \text{ if } k_2 \text{ even},$$
$$(x^J_{2k_1-1}, y^J_{k_2}) = (x^{J-1}_{k_1}, y^{J-1}_{k_2}), \quad k_1 = 1, ..., N_J/2 \text{ if } k_2 \text{ odd},$$

and for the second

$$(x^{J-1}_{k_1}, y^{J-1}_{2k_2}) = (x^{J-2}_{k_1}, y^{J-2}_{k_2}), \quad k_1, k_2 = 0, ..., N_J/2 - 1.$$

The following steps are performed similarly. Let us put $N_j := N_J/2^{(J-\lfloor j/2 \rfloor)}$. We consider $v^j_{k_1,k_2} = v(x^j_{k_1}, y^j_{k_2})$, where if j is even $0 \le k_1, k_2 \le N_j - 1$ and if j is odd $0 \le k_2 \le 2N_j$ and $1 \le k_1 \le N_j$. The projection operators are for j even $v^{j-1}_{k_1,k_2} = (P^j_{j-1}v^j)_{k_1,k_2} = v^j_{2k_1,k_2}$ if k_2 is even, $v^{j-1}_{k_1,k_2} = (P^j_{j-1}v^j)_{k_1,k_2} = v^j_{2k_1-1,k_2}$ if k_2 is odd, and for j odd $v^{j-1}_{k_1,k_2} = (P^j_{j-1}v^j)_{k_1,k_2} = v^j_{k_1,2k_2}$. In particular, we obtain for j even that the kernel of the linear operator of decimation, P^j_{j-1} , reads $Ker(P^j_{j-1}) = \{v^j \in V^j,\ v^j_{2k_1,k_2} = 0, k_2 \text{ even }, v^j_{2k_1-1,k_2} = 0, k_2 \text{ odd }\}$, and for j odd $Ker(P^j_{j-1}) = \{v^j \in V^j,\ v^j_{k_1,2k_2} = 0\}$. Thus, if we denote by e^k the prediction error, we will need to keep when j is even $e^j_{2k_1-1,k_2}$ for k_2 even, $e^j_{2k_1,k_2}$ for k_2 odd and $e^j_{k_1,2k_2-1}$ when j is odd. A reconstruction procedure for this discretization is given by

operator $S(v^{j-1}) := S(v^{j-1})v^{j-1}$ that interpolates the data v^{j-1} at the grid points of Γ^{j-1}. We consider three different types of prediction rules: linear prediction, ENO and WENO prediction. Let us remark that in the algorithm given below we denote $\hat{v}^j = v^j \setminus v^{j-1}$, that is, for all $j = 0, \ldots, J$, $v^j = \bigcup_{k=1}^{j} \hat{v}^k$. Then, encoding and decoding algorithms take the following form:

Encoding Algorithm

Quincunx Matrix

for j = J-2, . . . , 2; j = j - 2

for k_1,k_2 = 1, . . . , N_{j-1}

$\hat{v}^{j-2}_{k_1,k_2} = v^j_{2k_1-1,2k_2-1}$

$\hat{v}^{j-1}_{k_1,2k_2} = v^j_{2k_1,2k_2}$

end

for k_1, k_2=1, . . .,N_{j-1}

$e^{j-1}_{k_1,k_2} = v^j_{2k_1,2k_2} - S(v^{j-2})_{2k_1,2k_2}$

end

for $k_1 = 1, \ldots, N_{j-1}$

for k_2 = 1, . . . , $N_j; k_2 = k_2 + 2$

$e^j_{k_1,2k_2} = v^j_{2k_1,k_2} - S(v^{j-1})_{2k_1,k_2}$

$e^j_{k_1,k_2+1} = v^j_{2k_1-1,k_2+1} - S(v^{j-1})_{2k_1-1,k_2+1}$

end

end

end

Hexagonal Matrix

for j = J-2, . . . , 2; j = j - 2

for k_1,k_2 = 1, . . . , N_{j-1}

$\hat{v}^{j-2}_{k_1,k_2} = v^j_{4k_1-3,4k_2-3}$, $\hat{v}^{j-1}_{k_1,3k_2-2} = v^j_{4k_1-1,4k_2-2}$

$\hat{v}^{j-1}_{k_1,3k_2-1} = v^j_{4k_1-3,4k_2-1}$, $\hat{v}^{j-1}_{k_1,3k_2} = v^j_{4k_1-1,4k_2}$

end

for k_1, k_2=1, . . .,N_{j-1}

$e^{j-1}_{k_1,3k_2-2} = v^j_{4k_1-1,4k_2-2} - S(v^{j-2})_{4k_1-1,4k_2-2}$

$e^{j-1}_{k_1,3k_2-1} = v^j_{4k_1-3,4k_2-1} - S(v^{j-2})_{4k_1-3,4k_2-1}$

$e^{j-1}_{k_1,3k_2} = v^j_{4k_1-1,4k_2} - S(v^{j-2})_{4k_1-1,4k_2}$

end

for $k_1 = 1, \ldots, N_{j-1}$

for k_2 = 1, . . . , $N_j; k_2 = k_2 + 2$

$e^j_{3k_1-2,k_2} = v^j_{4k_1-2,k_2} - S(v^{j-1})_{4k_1-2,k_2}$

$e^j_{3k_1-1,k_2} = v^j_{4k_1-1,k_2} - S(v^{j-1})_{4k_1-1,k_2}$

$e^j_{3k_1,k_2} = v^j_{4k_1,k_2} - S(v^{j-1})_{4k_1,k_2}$

$e^j_{3k_1-2,k_2+1} = v^j_{4k_1-3,k_2+1} - S(v^{j-1})_{4k_1-3,k_2+1}$

$e^j_{3k_1-1,k_2+1} = v^j_{4k_1-3,k_2+1} - S(v^{j-1})_{4k_1-3,k_2+1}$

$e^j_{3k_1,k_2+1} = v^j_{4k_1,k_2+1} - S(v^{j-1})_{4k_1,k_2+1}$

end

end

end

The encoding algorithm leads to the following representation: $\mathcal{M}v^J = (v^0, d^0, ..., d^{J-1})$, where d^j was defined at the end of section III. Making the same reasoning, we can encode the multiresolution decomposition obtained with the hexagonal sampling matrix.

Then, we concatenate matrices $v^0, d^0, \ldots, d^{J-1}$ in one matrix V of dimension N^J (that is, of the same

dimension as v^J). To build V, we start by writing down matrix v^0 in its left upper corner. Then, we write down the matrix d^1 to the right of matrix v^0 and then d^1 below matrices v^0 and d^0 thus getting a square matrix. We proceed in the same way (writing down matrix d^j to the right of $d^j - 2$ if j is even and below if j is odd) until $j = J$. The result of this concatenation is represented on Figure 3.(A) for the Quincunx matrix and in Figure 3.(B). for the Hexagonal matrix (the subspace d^j is represented by D^j).

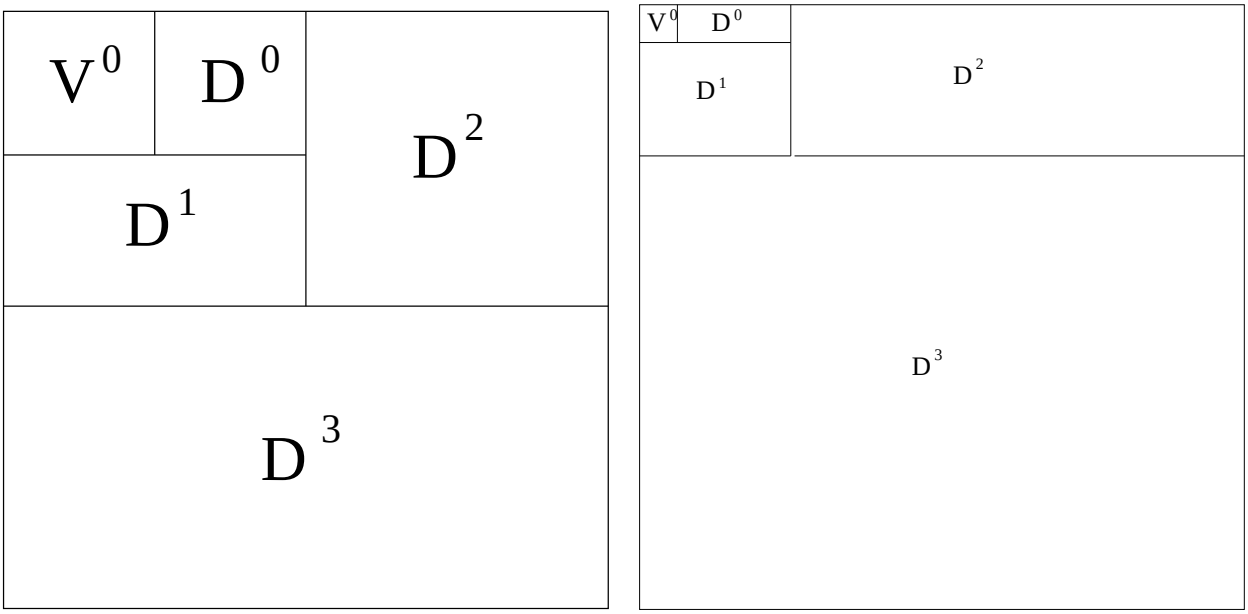


Fig. 3. (A): quincunx matrix, (B): hexagonal matrix.

Therefore, we are led to encode V. Let us notice that in general the prediction error is larger when the resolution is low. That is, the smaller j, the larger the elements of sub-matrix D^j. This naturally leads us to build an EZW (Embedded zero-tree wavelet) encoder adapted to our context.

The EZW (Embedded Zero-tree Wavelet) encoding algorithm is an encoder specially developed to use with wavelet transforms. It is based on progressive encoding: the data is compressed in multiple passes with increasing accuracy. The initial threshold is set to be equal to $T_0 = 2^{\lfloor \log_2(\max |V(k_1,k_2)|) \rfloor}$, where $\max |V(k_1, k_2)|$ means the maximum element of matrix V. The encoder scans next element, compares it with threshold and gives 'p', 'n', 'z' or 't' as an output; if the absolute value is bigger than the threshold, it outputs either 'p', if the value is positive or 'n', if it is negative, else it constructs a tree with the considered element as the root. If it is a zero-tree (that is, the values in the nodes are smaller or equal to the root), the output is 't', otherwise it is 'z' (isolated zero). We encode an element only if we got 'p'

or 'n' as output. In this case, we put it in the so-called 'subordinate list' together with either $\frac{3T_0}{!2}$, if it is bigger than T_0 or $-\frac{3T_0}{2}$ (for elements inferior to $-\frac{T_0}{2}$) and remove it from V (replace it by 0) so that it will not be encoded again. After we have scanned all the elements, we divide the threshold T_0 by two and start again the procedure till the threshold is not smaller than a fixed value.

Thereby, to apply an encoder based on the same principle as EZW encoder, we need to introduce a tree structure for them. The trees are designed in such a way that every element of the sub-matrix D^j is a child of some element of sub-matrix D^{j-1}, $j = 1, \ldots, 2\log_2 N_J$ for quincunx matrix, $j = 1, \ldots, 2\log_4 N_J$ for hexagonal matrix. That is, we get a binary tree for quincunx matrix and a quad-tree for hexagonal matrix. The tree structures for quincunx and hexagonal trees are illustrated on Figure 4.(A) and (B).

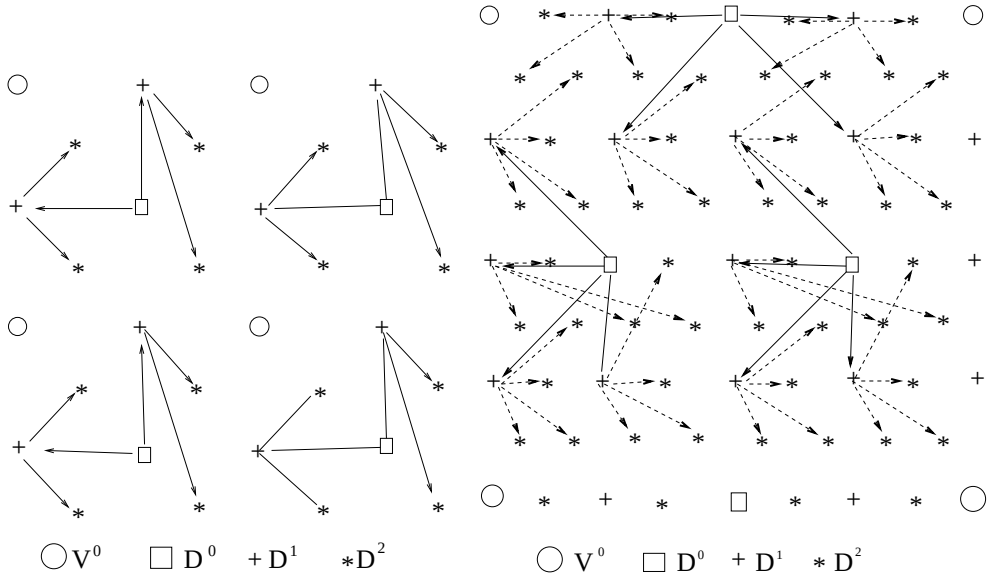


Fig. 4. (A): Tree structure for quincunx matrix, (B): Tree structure for hexagonal matrix.

In order that the algorithm works correctly, it is necessary that matrix V be scanned in such a way that sub-matrix D^j is scanned before D^{j+1}, subspace D^j being itself scanned using the so-called Morton scan. For each subspace D^j, we will denote by T_1^j (resp. T_2^j) the number of lines (resp. columns) in the matrix D^j. The EZW decoding is realized exactly in the same way as for wavelets. We now write the inverse operator $v^J = \mathcal{M}^{-1}(v^0, d^0, ..., d^{J-1})$.

Decoding Algorithm

Quincunx Matrix

for j = 2, . . . , J-1; j = j + 2

for k_1, k_2 = 1, . . . , N_{j-1}

$\hat{v}^{j-1}_{2k_1-1,2k_2-1} = S(v^{j-2})_{2k_1-1,2k_2-1}$

$\hat{v}^{j-1}_{2k_1,2k_2} = e^{j-1}_{k_1,k_2} + S(v^{j-2})_{2k_1,2k_2}$

end

for k_1=1, . . . , N_{j-1}

for k_2=1, . . . ,N_j; $k_2 = k_2 + 2$

$v^j_{2k_1,k_2} = e^j_{k_1,k_2} + S(v^{j-1})_{2k_1,k_2}$

$v^j_{2k_1-1,k_2} = e^j_{k_1,k_2+1} + S(v^{j-1})_{2k_1-1,k_2+1}$

end

end

Hexagonal Matrix

for j = 2, . . . , J-1; j = j + 2

for k_1, k_2 = 1, . . . , N_{j-1}

$\hat{v}^{j-1}_{4k_1-3,4k_2-3} = v^{j-2}_{k_1,k_2}$

$\hat{v}^{j-1}_{4k_1-1,4k_2-2} = e^{j-1}_{k_1,3k_2-2} + S(v^{j-2})_{4k_1-1,4k_2-2}$

$\hat{v}^{j-1}_{4k_1-3,4k_2-1} = e^{j-1}_{k_1,3k_2-1} + S(v^{j-2})_{4k_1-3,4k_2-1}$

$\hat{v}^{j-1}_{4k_1-1,4k_2} = e^{j-1}_{k_1,3k_2} + S(v^{j-2})_{4k_1-1,4k_2}$

end

for k_1=1, . . . , N_{j-1}

for k_2=1, . . . ,N_j; $k_2 = k_2 + 2$

$v^j_{4k_1-1,k_2} = e^j_{3k_1-1,k_2} + S(v^{j-1})_{4k_1-1,k_2}$

$v^j_{4k_1,k_2} = e^j_{3k_1,k_2} + S(v^{j-1})_{4k_1,k_2}$

$v^j_{4k_1,k_2} = e^j_{3k_1,k_2} + S(v^{j-1})_{4k_1,k_2}$

$v^j_{4k_1-3,k_2+1} = e^j_{3k_1-2,k_2+1} + S(v^{j-1})_{4k_1-3,k_2+1}$

$v^j_{4k_1-2,k_2+1} = e^j_{3k_1-1,k_2+1} + S(v^{j-1})_{4k_1-2,k_2+1}$

$v^j_{4k_1,k_2+1} = e^j_{3k_1,k_2+1} + S(v^{j-1})_{4k_1,k_2+1}$

end

end

VII. Numerical Applications

In this section, we study the improvement brought about either by the use of nonlinear affine instead of linear affine interpolant or by using nonlinear quadratic interpolant instead of a linear one. In particular, we will put forward the importance of dealing nonlinearly with the finest scales especially for geometric images. We carry out the study both for the quincunx and hexagonal dilation matrix. We also investigate the importance of the degree of the interpolation polynomial in terms of compression results.

A. Interpolation using Affine Polynomials

We investigated the improvement brought about the use of affine nonlinear prediction schemes instead of linear ones on natural images of Figure 5. In particular, we investigated the importance of nonlinear prediction at fine scales. To do so, we considered nonlinear prediction when T_1^j, the first dimension of D^j in the matrix V, is above T_1 and linear prediction elsewhere. For a 256×256 image and when the quincunx matrix is used, $T_1 = 64$ means that we predict nonlinearly the last finest four subspaces. We

made the simulation for $T_1 = 32, 64, 128$ for the 256×256 images of Lena (Figure 5.(A)) and of peppers (Figure 5.(B)). We also implemented the WENO prediction method.

The results we depict on Figure 6 (A) and (B) corresponds to the prediction rules (8) and (9). The cells on which we shall predict nonlinearly either corresponds to the cells where the average first order differences are locally maximum in the vertical or the horizontal directions (C_1 on Figure 6 (A) and (B), corresponding to rule (10) or the first order differences are locally maximum in the direction of prediction (C_2 on Figure 6 (A) and (B), corresponding to rule (11)).

When the quincunx matrix is used, a significant compression gain is observed when the nonlinear method is used especially at high compression rate (see Figure 6 (A) and (B)). A more careful look shows that one shall predict non linearly a larger number of scales when the image is geometric such as the image of Lena while the number of scales predicted nonlinearly are less importance for natural images such as the image of Lena. Furthermore, we noticed that when one uses the rule (11) instead of (10) to determine the points where the prediction is nonlinear, the gain in compression is significant for both images and even more for the image of peppers.

The results on WENO prediction are displayed on Figure 6 (C) and (D). For both images, the effect of using the WENO method is to make the prediction operator closer to the linear model which has the consequence that the compression gain over the linear model is significantly less important than when the ENO method is used. WENO method are important to construct stable multi-scale representation. We are currently looking for a different way to define the weights so as to preserve compression performance.

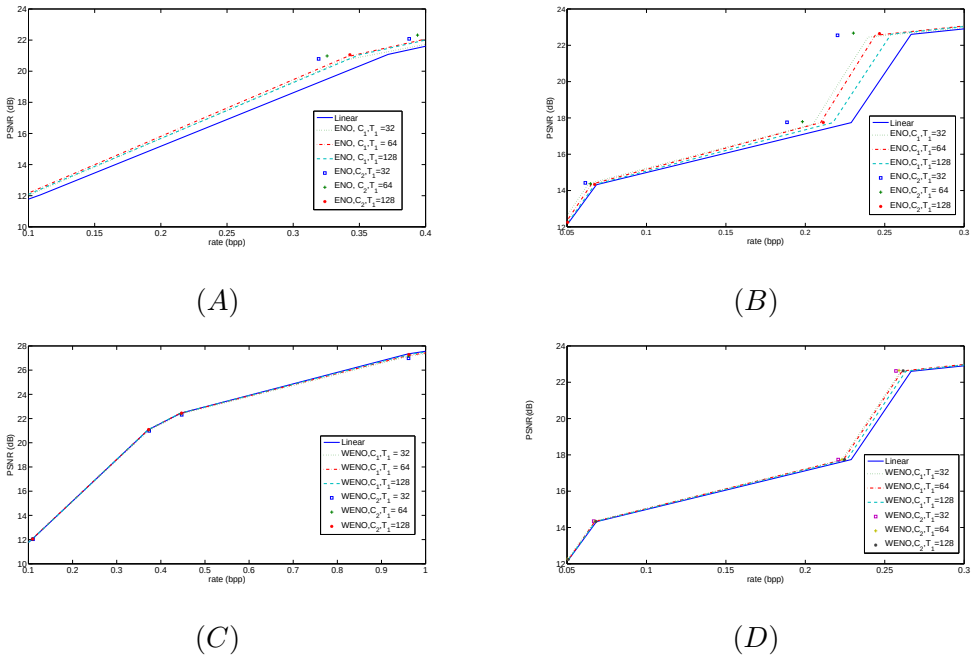
Now, if we switch on to the Hexagonal dilation matrix, the results are displayed on Figure (7).(C), we notice that the nonlinear prediction does not bring any improvement in terms of compression results (note that due to the greater down-sampling we considered $T_1 = 16$ or 64 which respectively corresponds to the last four subspaces are predicted nonlinearly or only the last 2). Our insight into these results is that the compression performance is related to $m = |\det(M)|$. Indeed, with the hexagonal matrix the sampling factor is too important and too much information is lost between each scale for the prediction to be efficient. Another possible direction for future research would be the optimization of the cost function that determines the stencil to use.

B. The Quadratic Prediction

We investigated the improvement brought about nonlinear quadratic prediction instead of linear quadratic prediction. The improvement in terms of compression rate occur essentially at intermediate compression rate as shown on Figure 7 (A) and (B) whereas with the affine prediction the gain was more important at

(A) (B)

Fig. 5. (A): a 256×256 Lena image, (B): a 256×256 peppers image



(A) (B)

(C) (D)

Fig. 6. (A): linear prediction (solid line) and ENO prediction for varying T_1 using either C_1 or C_2 to compute cells of interest for Lena image and when the quincunx matrix is used, (B):idem but for the image of peppers (C): linear prediction (solid line) and WENO predictions for varying T_1 using either C_1 or C_2 to compute cells of interest for Lena image and when the quincunx matrix is used, (D): the same as (C) but for peppers image

high compression rate. We notice again that it is less important to predict nonlinearly more scales with the image of Lena than with the image of peppers for which it appears crucial to predict a sufficient

number of scales to get better results than in the linear case (see Figure 7 (B)) .

Finally, when one uses the hexagonal matrix, Figure 7 (D) shows that to use higher degree polynomial does not bring any improvement in terms of compression results. In terms of perspective, future work should involve the definition of a cost function to determine which stencil to use that shall depend on the matrix M. To conclude the results section, we obtain significant improvement when the quincunx matrix is used while the hexagonal dilation matrix does not seem a good choice in our framework. We are currently trying to understand why the dilation matrix plays such an important role.

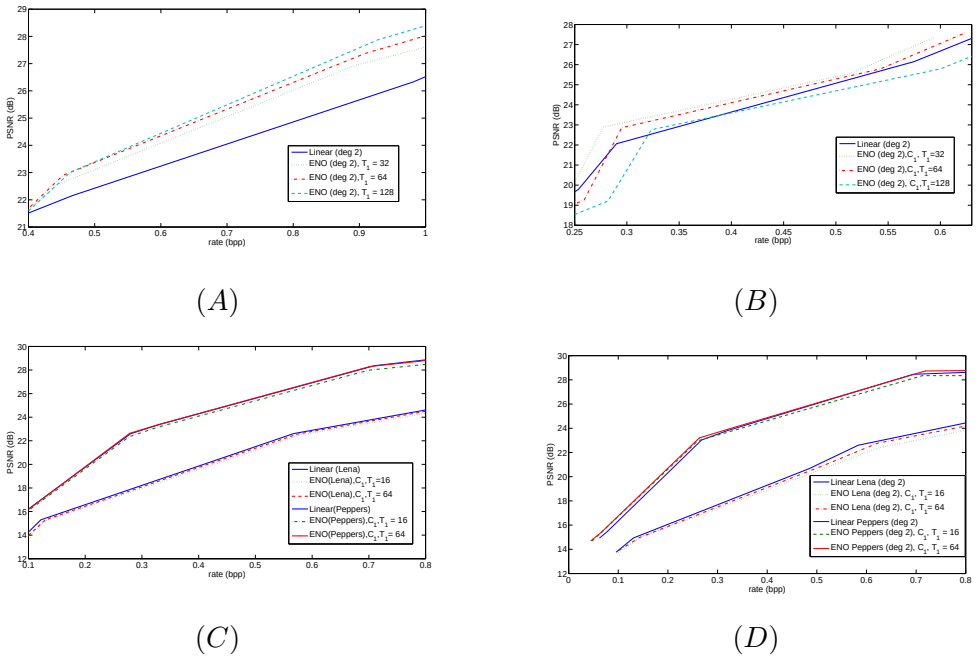


Fig. 7. (A): The quadratic linear prediction (solid line) is compared to the nonlinear quadratic prediction when the value of T_1 varies for Lena image and when the quincunx matrix is used, (B): idem but for the image of peppers, (C): Affine linear prediction (solid line) is compared to the nonlinear affine ENO prediction when the value of T_1 varies for both the image of Lena and the image of peppers and when the hexagonal matrix is used ,(D): Same computations as in (C) except that we use quadratic polynomial

VIII. Conclusion

In this paper, we have presented the theory of nonlinear and non-separable multi-scale representation. We have first emphasized the relation between nonlinear prediction operators and multi-scale representations. We have then shown that the convergence of the multiscale representation is related to some property of the joint spectral radius of the first order difference operator computed from the prediction operator. We

have then built some bidimensional nonlinear multi-scale representations based on interpolatory prediction operators. These were built using either the quincunx or the hexagonal matrix as dilation matrix and were either affine predictors or based on polynomials of degree 2. To show the pertinence of the approach we have proposed an application to image compression. To do so, we have first explained how the embedded-zero-tree wavelet (EZW) algorithm adapts in our context. The compression results show a clear improvement brought about ENO methods when the quincunx matrix is used both with affine or quadratic predictors. However, when the WENO prediction is used, the compression results is very close to the linear one, which let us think that future work should involve a deeper study of the weights in the WENO method. When the hexagonal matrix is used, we did not manage to obtain similar results as with the quincunx case ; we believe that a the down-sampling factor (symbolized in our context by the determinant of the matrix) is the key point, but this still needs further study. In terms of perspectives, we also are currently investigating potential extension of our approach to non-interpolatory prediction operators.

APPENDIX A

PROOF :Let $v, \tilde{v} \in \ell^\infty(\mathbb{Z}^2)$. By the definition of WENO prediction we have that

$$\|a_{k-Ml}(v) - a_{k-Ml}(\tilde{v})\|_{\ell^\infty(\mathbb{Z}^2)} \leq \|S(v) - S(\tilde{v})\|_{\ell^\infty(\mathbb{Z}^2)} \leq |\alpha_1 - \tilde{\alpha}_1| + |\alpha_2 - \tilde{\alpha}_2|,$$

where $\alpha_1, \tilde{\alpha}_1, \alpha_2, \tilde{\alpha}_2$ are the weights of the stencils for v and $\tilde{v}$ respectively. For $i = 1, 2$ it holds that

$$\begin{aligned}
|\alpha_i - \tilde{\alpha}_i| &= |\frac{a_i}{a_1 + a_2} - \frac{\tilde{a}_i}{\tilde{a}_1 + \tilde{a}_2}| \\
&\leq |\frac{a_i - \tilde{a}_i}{a_1 + a_2}| + |\tilde{a}_i(\frac{1}{a_1 + a_2} - \frac{1}{\tilde{a}_1 + \tilde{a}_2})| \\
&\leq |\frac{1}{a_1 + a_2}|(|a_i - \tilde{a}_i| + \frac{\tilde{a}_i}{\tilde{a}_1 + \tilde{a}_2}|a_1 - \tilde{a}_1 + a_2 - \tilde{a}_2| \\
&\leq |\frac{1}{a_1 + a_2}|2\sum_{i=1}^{2}|a_i - \tilde{a}_i|,
\end{aligned}$$

then, $\|S(v) - S(\tilde{v})\|_{\ell^\infty(\mathbb{Z}^2)} \leq \frac{4}{a_1+a_2}\sum_{i=1}^{2}|a_i - \tilde{a}_i|$. By simple computations we obtain that: $|a_i - \tilde{a}_i| = |\frac{1}{\epsilon+b_i} - \frac{1}{\epsilon+\tilde{b}_i}| \leq \frac{|b_i - \tilde{b}_i|}{\epsilon^2}$.

From the definition of b_i we have that $b_i \lesssim \|v\|_{\ell^\infty(\mathbb{Z}^2)}$ and $|b_i - \tilde{b}_i| \leq \|v - \tilde{v}\|_{\ell^\infty(\mathbb{Z}^2)}$ and it follows that $a_1 + a_2 = \frac{1}{\epsilon+b_1} + \frac{1}{\epsilon+b_1} \geq \frac{2}{\epsilon+\|v\|_{\ell^\infty(\mathbb{Z}^2)}}$. Therefore,

$$\sum_i |a_i - \tilde{a}_i| \lesssim \frac{\|v - \tilde{v}\|_{\ell^\infty(\mathbb{Z}^2)}}{\epsilon^2}$$

and $\|S(v) - S(\tilde{v})\|_{\ell^\infty(\mathbb{Z}^2)} \lesssim \frac{\epsilon+\|v\|_{\ell^\infty}}{\epsilon^2}\|v - \tilde{v}\|_{\ell^\infty(\mathbb{Z}^2)}$.

APPENDIX B

Let us consider first that $k = Mk'$ and then we compute (notice that $\varepsilon_1 = e_2$):

$$\begin{aligned} v^{j,1}_{Mk+\varepsilon_1} - v^j_{Mk} &= \frac{1}{2}(v^{j-1}_k + v^{j-1}_{k+e_1+e_2}) - v^{j-1}_k = \frac{1}{2}(v^{j-1}_{Mk'+e_1+e_2} - v^{j-1}_{Mk'}) \\ &= \frac{1}{2}(v^{j-2}_{k'+e_2} - v^{j-2}_{k'}) \end{aligned}$$

$$v^{j,2}_{Mk+\varepsilon_1} - v^j_{Mk} = \frac{1}{2}(v^{j-1}_{Mk'+e_1} + v^{j-1}_{Mk'+e_2}) - v^{j-1}_{Mk'} = \frac{1}{2}(v^{j-1}_{M(k'-e_2)+e_2} + v^{j-1}_{Mk'+e_2}) - v^{j-2}_{k'}$$

we then have to consider the different prediction rules for the values $v^{j-1}_{M(k'-e_2)+e_2}$ and $v^{j-1}_{Mk'+e_2}$. If the rule (8) is used in both cases, we have:

$$v^{j,1}_{Mk+\varepsilon_1} - v^j_{Mk} = \frac{1}{4}(v^{j-2}_{k-e_2} - v^{j-2}_k) + \frac{1}{2}(v^{j-2}_{k+e_1} - v^{j-2}_k) + \frac{1}{4}(v^{j-2}_{k+e_1+e_2} - v^{j-2}_{k+e_1}).$$

If the prediction rule (9) is used in each case, we have:

$$v^{j,1}_{Mk+\varepsilon_1} - v^j_{Mk} = \frac{1}{4}(v^{j-2}_{k-e_2+e_1} - v^{j-2}_{k-e_2}) + \frac{1}{4}(v^{j-2}_{k-e_2} - v^{j-2}_k) + \frac{1}{4}(v^{j-2}_{k+e_1} - v^{j-2}_k) + \frac{1}{4}(v^{j-2}_{k+e_2} - v^{j-2}_k).$$

If prediction (8) is used for $v^{j-1}_{M(k'-e_2)+e_2}$ and prediction (9) for $v^{j-1}_{Mk'+e_2}$, we may write:

$$v^{j,1}_{Mk+\varepsilon_1} - v^j_{Mk} = \frac{1}{2}(v^{j-2}_{k+e_1} - v^{j-2}_k) + \frac{1}{4}(v^{j-2}_{k-e_2} - v^{j-2}_k) + \frac{1}{4}(v^{j-2}_{k+e_2} - v^{j-2}_k).$$

If prediction (9) is used to predict $v^{j-1}_{M(k'-e_2)+e_2}$ and prediction (8) for $v^{j-1}_{Mk'+e_2}$, we may write:

$$v^{j,1}_{Mk+\varepsilon_1} - v^j_{Mk} = \frac{1}{4}(v^{j-2}_{k-e_2+e_1} - v^{j-2}_{k+e_1}) + \frac{1}{4}(v^{j-2}_{k+e_1} - v^{j-2}_k) + \frac{1}{4}(v^{j-2}_{k+e_1+e_2} - v^{j-2}_{k+e_1})$$

From the above equality, we can deduce that when $k = Mk'$:

$$\begin{aligned} \|v^{j,1}_{M.+\varepsilon_1} - v^j_{M.}\|_{\ell^\infty(\mathbb{Z}^2)} &\leq \frac{1}{2}\|\nabla v^{j-2}\|_{(l^\infty(\mathbb{Z}^2))^2} \\ \|v^{j,2}_{M.+\varepsilon_1} - v^j_{M.}\|_{\ell^\infty(\mathbb{Z}^2)} &\leq \|\nabla v^{j-2}\|_{(l^\infty(\mathbb{Z}^2))^2} \end{aligned}$$

Now, let us consider the case where $k = Mk' + \varepsilon_1$. If prediction (8) is used we may write:

$$\begin{aligned} v^{j,1}_{Mk+\varepsilon_1} - v^j_{Mk} &= \frac{1}{2}(v^{j-1}_{Mk'+e_2} + v^{j-1}_{M(k'+e_2)+e_2}) - v^{j-1}_{Mk'+e_2}) \\ &= \frac{1}{2}(v^{j-1}_{M(k'+e_2)+e_2} - v^{j-1}_{Mk'+e_2}) \end{aligned}$$

This corresponds to the same situation as that previously studied, and similar computation leads to:

$$\|v^{j,1}_{M.+\varepsilon_1} - v^j_{M.}\|_{\ell^\infty(\mathbb{Z}^2)} \leq \|\nabla v^{j-2}\|_{(\ell^\infty(\mathbb{Z}^2))^2}$$

If prediction (9) is used, we may write:

$$v^{j,2}_{Mk+\varepsilon_1} - v^j_{Mk} \;=\; \frac{1}{2}(v^{j-2}_{k'+e_2} + v^{j-2}_{k'+e_1+e_2}) - v^{j-1}_{Mk'+\varepsilon_1}$$

If prediction (8) is used to predict $v^{j-1}_{Mk'+\varepsilon_1}$, then we have:

$$v^{j,2}_{Mk+\varepsilon_1} - v^j_{Mk} \;=\; \frac{1}{2}(v^{j-2}_{k'+e_2} - v^{j-2}_{k'}),$$

while when prediction (9) is used, we get:

$$v^{j,2}_{Mk+\varepsilon_1} - v^j_{Mk} \;=\; \frac{1}{2}(v^{j-2}_{k'+e_1+e_2} - v^{j-2}_{k'+e_2}).$$

From this, we deduce that:

$$\|v^{j,2}_{M.+\varepsilon_1} - v^j_{M.}\|_{\ell^\infty(\mathbb{Z}^2)} \leq \frac{1}{2}\|\nabla v^{j-2}\|_{(\ell^\infty(\mathbb{Z}^2))^2},$$

This ends the proof.

APPENDIX C

PROOF: for the sake of simplicity, we will put ∇_i for ∇_{e_i}.

- If we use V^1_k and W^1_k for prediction, then we have the following behavior for the differences:
  1) $l = Mk$, $\nabla_1 v^j_{Mk} = \frac{1}{2}\nabla_1 v^{j-1}_k \quad \nabla_2 v^j_{Mk} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_2}$
  2) $l = Mk+\varepsilon_1$, $\nabla_1 v^j_{Mk+\varepsilon_1} = \frac{1}{2}\nabla_1 v^{j-1}_{k+e_1} \quad \nabla_2 v^j_{Mk+\varepsilon_1} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1}$
  3) $l = Mk+\varepsilon_2$, $\nabla_1 v^j_{Mk+\varepsilon_2} = \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_2} \quad \nabla_2 v^j_{Mk+\varepsilon_2} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1}$
  4) $l = Mk+\varepsilon_3$, $\nabla_1 v^j_{Mk+\varepsilon_3} = \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1+e_2} \quad \nabla_2 v^j_{Mk+\varepsilon_3} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_1+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1+e_2}$
- When the stencils V^2_k and W^2_k are used for the prediction, we obtain:
  1) $l = Mk$, $\nabla_1 v^j_{Mk} = \frac{1}{2}\nabla_1 v^{j-1}_k \quad \nabla_2 v^j_{Mk} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_2}$
  2) $l = Mk+\varepsilon_1$, $\nabla_1 v^j_{Mk+\varepsilon_1} = \frac{1}{2}\nabla_1 v^{j-1}_{k+e_1} \quad \nabla_2 v^j_{Mk+\varepsilon_1} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_1+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1+e_2}$
  3) $l = Mk+\varepsilon_2$, $\nabla_1 v^j_{Mk+\varepsilon_2} = \frac{1}{4}\nabla_1 v^{j-1}_k + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1+e_2} \quad \nabla_2 v^j_{Mk+\varepsilon_2} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1+e_2}$
  4) $l = Mk+\varepsilon_3$, $\nabla_1 v^j_{Mk+\varepsilon_3} = \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1+e_2} \quad \nabla_2 v^j_{Mk+\varepsilon_3} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1}$
- When the stencils V^1_k and W^2_k are used for the prediction
  1) $l = Mk$, $\nabla_1 v^j_{Mk} = \frac{1}{2}\nabla_1 v^{j-1}_k \quad \nabla_2 v^j_{Mk} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_k$
  2) $l = Mk+\varepsilon_1$, $\nabla_1 v^j_{Mk+\varepsilon_1} = \frac{1}{2}\nabla_1 v^{j-1}_{k+e_1} \quad \nabla_2 v^j_{Mk+\varepsilon_1} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1}$
  3) $l = Mk+\varepsilon_2$, $\nabla_1 v^j_{Mk+\varepsilon_2} = \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_2} \quad \nabla_2 v^j_{Mk+\varepsilon_2} = \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1} + \frac{1}{4}\nabla_1 v^{j-1}_k$
  4) $l = Mk+\varepsilon_3$ $\nabla_1 v^j_{Mk+\varepsilon_3} = \frac{1}{2}\nabla_1 v^{j-1}_{k+e_1+e_2} \quad \nabla_2 v^j_{Mk+\varepsilon_3} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1}$

- When the stencils V_2^k and W_1^k are used for the prediction

  1) $l = Mk$, $\nabla_1 v^j_{Mk} = \frac{1}{2}\nabla_1 v^{j-1}_k \quad \nabla_2 v^j_{Mk} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_2}$
  2) $l = Mk+\varepsilon_1$, $\nabla_1 v^j_{Mk+\varepsilon_1} = \frac{1}{2}\nabla_1 v^{j-1}_{k+e_1} \quad \nabla_2 v^j_{Mk+\varepsilon_1} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_1+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1+e_2}$
  3) $l = Mk+\varepsilon_2$, $\nabla_1 v^j_{Mk+\varepsilon_2} = \frac{1}{4}\nabla_1 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1+e_2} \quad \nabla_2 v^j_{Mk+\varepsilon_2} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1+e_2}$
  4) $l = Mk+\varepsilon_3$, $\nabla_1 v^j_{Mk+\varepsilon_3} = \frac{1}{2}\nabla_1 v^{j-1}_{k+e_1} \quad \nabla_2 v^j_{Mk+\varepsilon_3} = -\frac{1}{2}\nabla_2 v^{j-1}_{k+e_1+e_2} + \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1+e_2}$

To complete the proof of the contractivity, we also have to compute

$$\begin{aligned} v^j_{Mk+\varepsilon_2+e_2} - v^{j,1}_{Mk+\varepsilon_2} &= \frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} - \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1} \\ v^j_{Mk+\varepsilon_2+e_2} - v^{j,2}_{Mk+\varepsilon_2} &= \frac{1}{2}\nabla_2 v^{j-1}_{k+e_2} - \frac{1}{4}\nabla_1 v^{j-1}_{k+e_1+e_2} \end{aligned}$$

and also,

$$v^j_{Mk+\varepsilon_1+e_1} - v^j_{Mk+\varepsilon_1} = \frac{1}{2}\nabla_1 v^{j-1}_{k+e_1}.$$

REFERENCES

[1] S. Amat,S. Busquier, J.C Trillo, *Nonlinear Harten's Multiresolution on the Quincunx Pyramid* J. Comput. Appl. Math., vol. 189, pp. 555-567, 2003.

[2] A. Cohen, N. Dyn and B. Matei, *Quasi-linear Subdivision Schemes with Applications to ENO Interpolation,* Appl. Comput. Harmon. Anal., vol. 15,pp. 89-116, 2003.

[3] T. F. Chan, H.M. Zhou, *ENO-Wavelet Transforms for Piecewise Smooth Functions,* SIAM J. Numer. Anal., vol. 40, pp. 1369-1404, 2002.

[4] V. Chappelier, C. Guillemot, *Oriented Wavelet Transform for Image Compression and Denoising,* IEEE Trans. Image Process., vol. 15, no. 10, pp. 2892-2903, 2006.

[5] A. Harten, *Discrete Multiresolution Analysis and Generalized Wavelets*, J. Appl. Num. Math., vol. 12 , pp. 153-193, 1993.

[6] A. Harten, B. Enquist, S. Osher and S. Chakravarthy, *Uniformly high order accurate essentially non-oscillatory schemes III*, J. Comput. Phys., vol. 71, pp. 231-303,1987.

[7] R.Q. Jia, Q. Jiang, S.L. Lee, *Convergence of Cascade Algorithms in Sobolev Spaces and Integrals of Wavelets* Numer. Math., vol 91, pp. 453-473, 2002.

[8] S. Amat,R; Donat, J. Liandrat, J.C Trillo *Analysis of a New Nonlinear Subdivision Scheme Application in Image Processing,* Foundations of Computational Mathematics, pp. 193-225, 2006.

[9] S. Amat, J. Liandrat *On the stability of the PPH Nonlinear Multiresolution* Applied and Computational Harmonic Analysis,vol. 18, pp. 198-206, 2005.

[10] B. Matei, *Smoothness Characterization and Stability in Nonlinear Multi-scale Framework: Theoretical Results,* Asymptotic Analysis, vol. 46, pp. 277-309, 2005.

[11] S. Amat, S. Busquier, J.C. Trillo, *Nonlinear Harten's Multiresolution on the Quincunx Pyramid,* Journal of Computational Mathematics, vol. 189, pp. 555-567, 2006.